

RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.**

Application Serial Number: 10/802,891

Source: _____

Date Processed by STIC: _____

ENTERED



IFWO

RAW SEQUENCE LISTING

DATE: 01/12/2005

PATENT APPLICATION: US/10/802,891

TIME: 15:07:45

Input Set : N:\Crf3\RULE60\10802891.raw.txt

Output Set: N:\CRF4\01122005\J802891.raw

SEQUENCE LISTING

```

3 (1) GENERAL INFORMATION:
5     (i) APPLICANT: Seidel, Christoph; Weinhues, Ursula-Henrike;
6             Schmitt, Urban; Motz, Manfred; Wiedmann, Michael;
7             Upmeier, Barbara; Soutscheck, Erwin
10    (ii) TITLE OF INVENTION: Recombinant antigen from the NS3 region of
11                             the hepatitis C virus
13    (iii) NUMBER OF SEQUENCES: 9
16    (iv) CORRESPONDENCE ADDRESS:
17            (A) ADDRESSEE: Felfe & Lynch
18            (B) STREET: 805 Third Avenue
19            (C) CITY: New York
20            (D) STATE: New York
21            (E) COUNTRY: USA
22            (F) ZIP: 10022
25    (v) COMPUTER READABLE FORM:
26            (A) MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
27            (B) COMPUTER: IBM PS/2
28            (C) OPERATING SYSTEM: PC-DOS
29            (D) SOFTWARE: PatentIn Release #1.0,
30                           Version #1.25 (EPA)
33    (vi) CURRENT APPLICATION DATA:
C--> 34            (A) APPLICATION NUMBER: US/10/802,891
C--> 35            (B) FILING DATE: 17-Mar-2004
36            (C) CLASSIFICATION: 435
38    (vii) PRIOR APPLICATION DATA:
W--> 39            (A) APPLICATION NUMBER: US/08/511,759
40            (B) FILING DATE: 7-AUGUST-1995
W--> 41            (A) APPLICATION NUMBER: p 44 28 705.4
42            (B) FILING DATE: 12 AUGUST 1994
44    (viii) ATTORNEY/AGENT INFORMATION:
45            (A) NAME: HANSON, NORMAN
46            (B) REGISTRATION NUMBER: 30, 946
47            (C) REFERENCE/DOCKET NUMBER: HUBR 1067
49    (ix) TELECOMMUNICATION INFORMATION:
50            (A) TELEPHONE: 212-688-9200
51            (B) TELEFAX: 212-838-3884
55 (2) INFORMATION FOR SEQ ID NO: 1:
57     (i) SEQUENCE CHARACTERISTICS:
58            (A) LENGTH: 885 base pairs
59            (B) TYPE: nucleic acid
60            (C) STRANDEDNESS: both
61            (D) TOPOLOGY: linear

```

RAW SEQUENCE LISTING

DATE: 01/12/2005

PATENT APPLICATION: US/10/802,891

TIME: 15:07:45

Input Set : N:\Crf3\RULE60\10802891.raw.txt

Output Set: N:\CRF4\01122005\J802891.raw

```

63      (ii) MOLECULE TYPE: cDNA
65      (vi) ORIGINAL SOURCE:
66          (A) ORGANISM: hepatitis C virus
68      (viii) POSITION IN GENOME:
69          (A) CHROMOSOME/SEGMENT: NS3
71      (ix) FEATURE:
72          (A) NAME/KEY: CDS
73          (B) LOCATION: 1..885
77      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
79 ATG ACC ATG ATT ACG AAT TCC CGG GGA TCC ATC ATG AAA TCC CCG GTG      48
80 Met Thr Met Ile Thr Asn Ser Arg Gly Ser Ile Met Lys Ser Pro Val
81   1           5           10           15
83 TTC ACG GAT AAC TCC TCT CCA CCG GTA GTG CCC CAG AGC TTC CAG GTG      96
84 Phe Thr Asp Asn Ser Ser Pro Pro Val Pro Gln Ser Phe Gln Val
85           20           25           30
87 GCT CAC CTG CAT GCT CCC ACA GGC AGC GGC AAG AGC ACC AAG GTC CCG      144
88 Ala His Leu His Ala Pro Thr Gly Ser Gly Lys Ser Thr Lys Val Pro
89           35           40           45
91 GCT GCA TAC GCA GCT CAG GGC TAC AAG GTG CTA GTG CTC AAC CCT TCT      192
92 Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro Ser
93           50           55           60
95 GTT GCT GCA ACA TTG GGC TTT GGT GCC TAC ATG TCC AAG GCT CAT GGG      240
96 Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys Ala His Gly
97   65           70           75           80
99 ATC GAT CCT AAC ATC AGG ACC GGG GTG AGA ACA ATT ACC ACT GGC AGC      288
100 Ile Asp Pro Asn Ile Arg Thr Gly Val Arg Thr Ile Thr Thr Gly Ser
101           85           90           95
103 CCC ATT ACG TAC TCC ACT TAC GGC AAG TTT CTT GCC GAC GGC GGG TGC      336
104 Pro Ile Thr Tyr Ser Thr Tyr Gly Lys Phe Leu Ala Asp Gly Gly Cys
105           100          105          110
107 GCA GGG GGT GCT TAT GAC ATA ATA ATT TGT GAC GAG TGC CAC TCC ACG      384
108 Ala Gly Gly Ala Tyr Asp Ile Ile Cys Asp Glu Cys His Ser Thr
109           115          120          125
111 GAT GCC ACA TCC ATC TTG GGC ATC GGC ACT GTC CTT GAC CAA GGA GAG      432
112 Asp Ala Thr Ser Ile Leu Gly Ile Gly Thr Val Leu Asp Gln Gly Glu
113           130          135          140
115 ACT GCG GGG GCG AAA TTG GTT GTG TTC GCC ACC GCC ACC CCT CCG GGC      480
116 Thr Ala Gly Ala Lys Leu Val Val Phe Ala Thr Ala Thr Pro Pro Gly
117 145           150          155          160
119 TCC GTC ACT GTG CCC CAT CCC AAC ATT GAG GAG GTT GCT CTA TCC ACC      528
120 Ser Val Thr Val Pro His Pro Asn Ile Glu Glu Val Ala Leu Ser Thr
121           165          170          175
123 ACC GGA GAG ATC CCT TTT TAC GGC AAG GCT ATC CCC CTT GAG GTA ATC      576
124 Thr Gly Glu Ile Pro Phe Tyr Gly Lys Ala Ile Pro Leu Glu Val Ile
125           180          185          190
127 AAG GGG GGG AGA CAT CTC ATC TTC TGT CAT TCA AAG AGG AAG TGC GAT      624
128 Lys Gly Gly Arg His Leu Ile Phe Cys His Ser Lys Arg Lys Cys Asp
129           195          200          205
131 GAG CTC GCC ACA AAG CTG GTC GCA ATG GGC ATC AAT GCC GTG GCC TAC      672

```

RAW SEQUENCE LISTING

DATE: 01/12/2005

PATENT APPLICATION: US/10/802,891

TIME: 15:07:45

Input Set : N:\Crf3\RULE60\10802891.raw.txt

Output Set: N:\CRF4\01122005\J802891.raw

```

132 Glu Leu Ala Thr Lys Leu Val Ala Met Gly Ile Asn Ala Val Ala Tyr
133      210                215                220
135 TAC CGC GGT CTT GAC GTG TCC GTC ATC CCG ACC AGC GGT GAT GTT GTC      720
136 Tyr Arg Gly Leu Asp Val Ser Val Ile Pro Thr Ser Gly Asp Val Val
137 225                230                235                240
139 GTC GTG GCA ACC GAC GCC CTC ATG ACC GGC TAT ACC GGC GAC TTC GAC      768
140 Val Val Ala Thr Asp Ala Leu Met Thr Gly Tyr Thr Gly Asp Phe Asp
141                245                250                255
143 TCG GTG ATA GAC TGC AAC ACG TGT GTC ACT CAG ACA GTC GAT TTC AGC      816
144 Ser Val Ile Asp Cys Asn Thr Cys Val Thr Gln Thr Val Asp Phe Ser
145                260                265                270
147 CTT GAC CCT ACC TTC ACC ATT GAG ACG ACC ACA CTT CCC CAG GAT GCT      864
148 Leu Asp Pro Thr Phe Thr Ile Glu Thr Thr Thr Leu Pro Gln Asp Ala
149                275                280                285
151 GTC TCC CGC ACT CAA CGA CGG      885
152 Val Ser Arg Thr Gln Arg Arg
153      290                295
160 (2) INFORMATION FOR SEQ ID NO: 2:
162      (i) SEQUENCE CHARACTERISTICS:
163          (A) LENGTH: 295 amino acids
164          (B) TYPE: amino acid
165          (D) TOPOLOGY: linear
167      (ii) MOLECULE TYPE: protein
169      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
172 Met Thr Met Ile Thr Asn Ser Arg Gly Ser Ile Met Lys Ser Pro Val
173   1                5                10                15
175 Phe Thr Asp Asn Ser Ser Pro Pro Val Val Pro Gln Ser Phe Gln Val
176                20                25                30
178 Ala His Leu His Ala Pro Thr Gly Ser Gly Lys Ser Thr Lys Val Pro
179                35                40                45
181 Ala Ala Tyr Ala Ala Gln Gly Tyr Lys Val Leu Val Leu Asn Pro Ser
182                50                55                60
184 Val Ala Ala Thr Leu Gly Phe Gly Ala Tyr Met Ser Lys Ala His Gly
185   65                70                75                80
187 Ile Asp Pro Asn Ile Arg Thr Gly Val Arg Thr Ile Thr Thr Gly Ser
188                85                90                95
190 Pro Ile Thr Tyr Ser Thr Tyr Gly Lys Phe Leu Ala Asp Gly Gly Cys
191                100               105               110
193 Ala Gly Gly Ala Tyr Asp Ile Ile Ile Cys Asp Glu Cys His Ser Thr
194                115               120               125
196 Asp Ala Thr Ser Ile Leu Gly Ile Gly Thr Val Leu Asp Gln Gly Glu
197                130               135               140
199 Thr Ala Gly Ala Lys Leu Val Val Phe Ala Thr Ala Thr Pro Pro Gly
200 145                150                155                160
202 Ser Val Thr Val Pro His Pro Asn Ile Glu Glu Val Ala Leu Ser Thr
203                165                170                175
205 Thr Gly Glu Ile Pro Phe Tyr Gly Lys Ala Ile Pro Leu Glu Val Ile
206                180                185                190
208 Lys Gly Gly Arg His Leu Ile Phe Cys His Ser Lys Arg Lys Cys Asp

```

RAW SEQUENCE LISTING

DATE: 01/12/2005

PATENT APPLICATION: US/10/802,891

TIME: 15:07:45

Input Set : N:\Crf3\RULE60\10802891.raw.txt

Output Set: N:\CRF4\01122005\J802891.raw

```

209          195          200          205
211 Glu Leu Ala Thr Lys Leu Val Ala Met Gly Ile Asn Ala Val Ala Tyr
212          210          215          220
214 Tyr Arg Gly Leu Asp Val Ser Val Ile Pro Thr Ser Gly Asp Val Val
215 225          230          235          240
217 Val Val Ala Thr Asp Ala Leu Met Thr Gly Tyr Thr Gly Asp Phe Asp
218          245          250          255
220 Ser Val Ile Asp Cys Asn Thr Cys Val Thr Gln Thr Val Asp Phe Ser
221          260          265          270
223 Leu Asp Pro Thr Phe Thr Ile Glu Thr Thr Thr Leu Pro Gln Asp Ala
224          275          280          285
226 Val Ser Arg Thr Gln Arg Arg
227          290          295
231 (2) INFORMATION FOR SEQ ID NO: 3:
234 (i) SEQUENCE CHARACTERISTICS:
235 (A) LENGTH: 40 base pairs
236 (B) TYPE: nucleic acid
237 (C) STRANDEDNESS: single
238 (D) TOPOLOGY: linear
240 (ii) MOLECULE TYPE: cDNA
242 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
245 AAGGGATCCA TCATGAAATC CCCGGTGTTC ACGGATAACT 40
250 (2) INFORMATION FOR SEQ ID NO: 4:
253 (i) SEQUENCE CHARACTERISTICS:
254 (A) LENGTH: 39 base pairs
255 (B) TYPE: nucleic acid
256 (C) STRANDEDNESS: single
257 (D) TOPOLOGY: linear
259 (ii) MOLECULE TYPE: cDNA
261 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
263 GGGAAGCCTT AATTCTTACC GTCGTTGAGT GCGGGAGAC 39
265 (2) INFORMATION FOR SEQ ID NO: 5:
268 (i) SEQUENCE CHARACTERISTICS:
269 (A) LENGTH: 39 base pairs
270 (B) TYPE: nucleic acid
271 (C) STRANDEDNESS: single
272 (D) TOPOLOGY: linear
274 (ii) MOLECULE TYPE: cDNA
277 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
279 GAGGGATCCA TCATGAAAGC GGTGGACTTT ATCCCTGTG 39
282 (2) INFORMATION FOR SEQ ID NO: 6:
285 (i) SEQUENCE CHARACTERISTICS:
286 (A) LENGTH: 33 base pairs
287 (B) TYPE: nucleic acid
288 (C) STRANDEDNESS: single
289 (D) TOPOLOGY: linear
291 (ii) MOLECULE TYPE: cDNA
294 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
296 GAGAAGCTTT TAACACGTGT TGCAGTCTAT CAC 33

```

RAW SEQUENCE LISTING

DATE: 01/12/2005

PATENT APPLICATION: US/10/802,891

TIME: 15:07:45

Input Set : N:\Crf3\RULE60\10802891.raw.txt

Output Set: N:\CRF4\01122005\J802891.raw

298 (2) INFORMATION FOR SEQ ID NO: 7:

301 (i) SEQUENCE CHARACTERISTICS:

302 (A) LENGTH: 39 base pairs

303 (B) TYPE: nucleic acid

304 (C) STRANDEDNESS: single

305 (D) TOPOLOGY: linear

307 (ii) MOLECULE TYPE: cDNA

310 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

312 GAGGGATCCA TCATGAAACA CCTGCATGCT CCCACCGGC 39

315 (2) INFORMATION FOR SEQ ID NO: 8:

318 (i) SEQUENCE CHARACTERISTICS:

319 (A) LENGTH: 33 base pairs

320 (B) TYPE: nucleic acid

321 (C) STRANDEDNESS: single

322 (D) TOPOLOGY: linear

324 (ii) MOLECULE TYPE: cDNA

327 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

329 GAGAAGCTTT TAATACCAAG CACAGCCTGC GTC 33

333 (2) INFORMATION FOR SEQ ID NO: 9:

335 (i) SEQUENCE CHARACTERISTICS:

336 (A) LENGTH: 302 amino acids

337 (B) TYPE: amino acid

338 (D) TOPOLOGY: linear

340 (ii) MOLECULE TYPE: protein

342 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

345	Pro	Val	Glu	Asn	Leu	Glu	Thr	Thr	Met	Arg	Ser	Pro	Val	Phe	Thr
346	1			5					10					15	
348	Asp	Asn	Ser	Ser	Pro	Pro	Val	Val	Pro	Gln	Ser	Phe	Gln	Val	Ala
349				20					25					30	
351	His	Leu	His	Ala	Pro	Thr	Gly	Ser	Gly	Lys	Ser	Thr	Lys	Val	Pro
352				35					40					45	
354	Ala	Ala	Tyr	Ala	Ala	Gln	Gly	Tyr	Lys	Val	Leu	Val	Leu	Asn	Pro
355				50					55					60	
357	Ser	Val	Ala	Ala	Thr	Leu	Gly	Phe	Gly	Ala	Tyr	Met	Ser	Lys	Ala
358				65					70					75	
360	His	Gly	Ile	Asp	Pro	Asn	Ile	Arg	Thr	Gly	Val	Arg	Thr	Ile	Thr
361				80					85					90	
363	Thr	Gly	Ser	Pro	Ile	Thr	Tyr	Ser	Thr	Tyr	Gly	Lys	Phe	Leu	Ala
364				95					100					105	
368	Asp	Gly	Gly	Cys	Ala	Gly	Gly	Ala	Tyr	Asp	Ile	Ile	Ile	Cys	Asp
369				110					115					120	
371	Glu	Cys	His	Ser	Thr	Asp	Ala	Thr	Ser	Ile	Leu	Gly	Ile	Gly	Thr
372				125					130					135	
374	Val	Leu	Asp	Gln	Gly	Glu	Thr	Ala	Gly	Ala	Lys	Leu	Val	Val	Phe
375				140					145					150	
377	Ala	Thr	Ala	Thr	Pro	Pro	Gly	Ser	Val	Thr	Val	Pro	His	Pro	Asn
378				155					160					165	
380	Ile	Glu	Glu	Val	Ala	Leu	Ser	Thr	Thr	Gly	Glu	Ile	Pro	Phe	Tyr
381				170					175					180	

VERIFICATION SUMMARY

DATE: 01/12/2005

PATENT APPLICATION: US/10/802,891

TIME: 15:07:46

Input Set : N:\Cr3\RULE60\10802891.raw.txt

Output Set: N:\CRF4\01122005\J802891.raw

L:34 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:35 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]

L:41 M:238 W: Alpha Fields not Ordered, Reordered [(A) APPLICATION NUMBER:] of (1) (vii)